

# Representation Theorems for Strong Predicate Exchangeability in Pure Inductive Logic

Malte S. Kließ

malte@kappaplus.de

2nd July 2015

In Pure Inductive Logic, the principle of Strong Predicate Exchangeability is a rational principle based on symmetry that sits in between the principles of Predicate Exchangeability and Atom Exchangeability. We will show a de Finetti – style representation theorem for probability functions that satisfy this principle in addition to Unary Language Invariance.

Key words: Strong Predicate Exchangeability, Language Invariance, Inductive Logic, Probability Logic, Uncertain Reasoning.

## 1 Introduction

A recurring theme in the study of Pure Inductive Logic is symmetry. Amongst such rational principles based on symmetry are the well-known principles of Atom and Predicate Exchangeability<sup>1</sup>, respectively.

The rational principle of Strong Predicate Exchangeability arose from a generalized version of the  $u^{\bar{p}}$  functions (see [6, Chapter 29]) that are the building blocks<sup>2</sup> of functions satisfying the principles of Atom Exchangeability and Unary Language Invariance. In

---

<sup>1</sup>In the case of Predicate Exchangeability, we would like to point out that while the principle, in various formulations, was accepted as a basic condition for probability functions by both Rudolf Carnap and W.E. Johnson, the principle only recently was studied in more detail, providing representation theorems.

<sup>2</sup>By ‘building blocks’ we mean that any function satisfying some rational principle can be represented as a convex combination of the building block functions.

fact, these generalized functions turned out to be just too strong to use as building blocks for Unary Language Invariance.

With this in mind, and especially with the rather artificial motivating example above, one might think that Strong Predicate Exchangeability is a rather artificial principle and thus might only be of interest as a technical exercise. There is however an alternative way of obtaining Strong Predicate Exchangeability via a weak form of Johnson's Sufficientness Postulate (JSP): just as JSP gives rise to Carnap's Continuum of Inductive Methods whose members satisfy Atom Exchangeability, so can we obtain a collection of functions arising from a weaker form of JSP that are characterized by Strong Predicate Exchangeability (see [4][Chapter 3.3]). We would then argue that while there may be no compelling argument for a rational agent to adopt Strong Predicate Exchangeability as a principle in its own right, it would be a consequence for any agent picking this weaker form of JSP.

The aim of this paper is to present de Finetti – style representation theorems for probability functions satisfying Strong Predicate Exchangeability in conjunction with Unary Language Invariance. We will show that any probability function satisfying Strong Predicate Exchangeability and Unary Language Invariance can be represented as a convex combination of functions that in addition also satisfy Weak Irrelevance.

## 2 Notation and Principles

The context of this paper is unary Pure Inductive Logic. We will be concerned with first order languages that only contain finitely many predicate symbols  $P_i$  and countably many constant symbols  $a_1, a_2, \dots, a_n, \dots$  which we can think of as exhausting the universe. For  $k \in \mathbb{N}^+ = \{n \in \mathbb{N} \mid n > 0\}$ , let  $L_k$  be the language containing only the predicates  $P_1, P_2, \dots, P_k$ . Unless otherwise stated we will take our default language  $L$  to be the language  $L_q$ . Let  $QFSL$  denote the set of quantifier-free sentences of  $L$  and  $SL$  the set of sentences of  $L$ .

Let  $L = L_q$ . An *atom* of  $L$  is a formula

$$\alpha(x) = \bigwedge_{i=1}^q \pm P_i(x),$$

where  $+P_i(x)$ ,  $-P_i(x)$  stand for  $P_i(x)$ ,  $\neg P_i(x)$ , respectively. Note that for  $L = L_q$ , there are  $2^q$  atoms, denoted  $\alpha_1, \alpha_2, \dots, \alpha_{2^q}$ , which we will assume to be in the usual lexicographic ordering.<sup>3</sup>

---

<sup>3</sup>In the usual lexicographic ordering,  $\alpha_1 = \bigwedge_{i=1}^q P_i(x)$ ,  $\alpha_2 = \bigwedge_{i=1}^{q-1} P_i(x) \wedge \neg P_q(x)$ ,  $\alpha_3 = \bigwedge_{i=1}^{q-2} P_i(x) \wedge \neg P_{q-1}(x) \wedge P_q(x)$ , etc., up to  $\alpha_{2^q} = \bigwedge_{i=1}^q \neg P_i(x)$ .

A *state description* of  $L$  for  $a_1, \dots, a_n$  is a sentence

$$\Theta(a_1, \dots, a_n) = \bigwedge_{i=1}^n \alpha_{h_i}(a_i),$$

where  $\alpha_{h_i}$  is an  $L$ -atom for each  $h_i$ , i.e.  $h_i \in \{1, \dots, 2^q\}$  for each  $1 \leq i \leq n$ .

The following definition will turn out to be very useful in keeping track of the number of negations occurring in an atom.

**Definition 1:**

Let  $L = L_q$ . Define the function  $\gamma_q$  on the atoms of  $L$  by

$$\gamma_q(\alpha) = k \Leftrightarrow \alpha \text{ contains } k \text{ negated predicates.}$$

We shall drop the index  $q$  whenever it is understood from the context.

**Definition 2:**

Let  $L = L_q$  and let  $w : SL \rightarrow [0, 1]$ . Then  $w$  is a probability function if  $w$  satisfies the following properties.

- (i)  $w(\top) = 1$ ,
- (ii) if  $\varphi \models \neg\vartheta$ , then  $w(\varphi \vee \vartheta) = w(\varphi) \cdot w(\vartheta)$ ,
- (iii)  $w(\exists x \varphi(x)) = \lim_{n \rightarrow \infty} w(\bigvee_{i=1}^n \varphi(a_i))$ .

Recall that by Gaifman's Theorem (see [3]) any probability function  $w : QFSL \rightarrow [0, 1]$  satisfying (i) and (ii) above can uniquely be extended to a function  $w$  on  $SL$  satisfying (i)-(iii). By the Disjunct Normal Form Theorem it will be enough to define a function  $w$  on state descriptions of  $L$ .

In the following definitions for rational principles we will assume that  $L = L_q$  and  $w$  is a probability function on  $SL$ .

**Constant Exchangeability (Ex)**

$w$  satisfies Constant Exchangeability if for all  $L$ -sentences  $\varphi$  and all permutations  $\sigma$  of  $\mathbb{N}^+$ ,

$$w(\varphi(a_1, \dots, a_n)) = w(\varphi(a_{\sigma(1)}, \dots, a_{\sigma(n)})).$$

We will assume that all probability functions satisfy Constant Exchangeability.

**Predicate Exchangeability (Px)**

$w$  satisfies Px if for all  $L$ -sentences  $\varphi$  and all permutations  $\sigma$  of the (indices of) predicates of  $L$ ,

$$w(\varphi(P_1, \dots, P_m, a_1, \dots, a_n)) = w(\varphi(P_{\sigma(1)}, \dots, P_{\sigma(m)}, a_1, \dots, a_n)).$$

**Atom Exchangeability (Ax)**

$w$  satisfies Atom Exchangeability if for all permutations  $\sigma$  of the (indices of) atoms of  $L$ ,

$$w\left(\bigwedge_{i=1}^n \alpha_{h_i}(a_i)\right) = w\left(\bigwedge_{i=1}^n \alpha_{\sigma(h_i)}(a_i)\right).$$

**Weak Irrelevance Principle (WIP)**

$w$  satisfies Weak Irrelevance if whenever  $\vartheta, \varphi \in QFSL$  have no constants or predicates in common then

$$w(\vartheta \wedge \varphi) = w(\vartheta) \cdot w(\varphi).$$

**Unary Language Invariance (ULi)**

Let  $w$  be a probability function on  $L$ . Then  $w$  satisfies Unary Language Invariance if there is a family of probability functions  $w^{\mathcal{L}}$ , one for each finite unary language  $\mathcal{L}$  and satisfying Px + Ex such that  $w^L = w$  and whenever  $\mathcal{L} \subseteq \mathcal{L}'$ , then  $w^{\mathcal{L}'} \upharpoonright S\mathcal{L} = w^{\mathcal{L}}$ .

We say that  $w$  satisfies ULi *with*  $\mathcal{P}$  for a rational principle  $\mathcal{P}$ , if each  $w^{\mathcal{L}}$  in the ULi family also satisfy  $\mathcal{P}$ .

We will now proceed to define the  $P$ -spectrum and in turn the principle of Spectrum Exchangeability, which is the main focus of this paper.

**Definition 3 ( $P$ -spectrum):**

Let  $L = L_q$  and  $\Theta(a_1, \dots, a_n) = \bigwedge_{i=1}^n \alpha_{h_i}(a_i)$  be a state description of  $L$ . We can view  $\Theta$  as

$$\Theta(a_1, \dots, a_n) = \bigwedge_{i=0}^q \bigwedge_{j=1}^{s_i} \alpha_{h_{ij}}(b_{ij}),$$

where each  $\alpha_{h_{ij}}$  for  $j \in \{1, \dots, s_i\}$  is an atom of  $L$  with  $\gamma_q(h_{ij}) = i$ ,  $n = \sum_{i=0}^q s_i$  and  $b_{ij} = a_k$  for that  $k$  with  $\Theta(a_1, \dots, a_n) \models \alpha_{h_{ij}}(a_k)$ . For each  $i \in \{0, \dots, q\}$  let  $E_i$  be the equivalence relation on  $\{1, \dots, s_i\}$  given by

$$jE_i k \iff h_{ij} = h_{ik}.$$

For  $i \in \{0, \dots, q\}$  let  $M_i$  be the multiset of sizes of equivalence classes of  $E_i$ . The  $P$ -spectrum of  $\Theta$  is the vector

$$\langle M_0, M_1, \dots, M_q \rangle.$$

**Definition 4 (Strong Predicate Exchangeability, SPx):**

Let  $L = L_q$  and  $w$  a probability function on  $L$ .  $w$  satisfies Strong Predicate Exchangeability if and only if  $w(\Theta) = w(\Phi)$  whenever  $\Theta$  and  $\Phi$  have the same  $P$ -spectrum.

SPx is a strong version of Px in the sense that it implies Px, but Px does not imply SPx.

**Proposition 5.** *Let  $w$  be a probability function on  $L = L_q$  satisfying SPx. Then  $w$  satisfies Px.*

**Proof:** Let  $\sigma$  be a permutation of predicates,  $\Theta$  a state description of  $L$  and  $\sigma\Theta$  the result of permuting all predicates occurring in  $\Theta$  according to  $\sigma$ . It is enough to show that  $\Theta$  and  $\sigma\Theta$  have the same  $P$ -spectrum.

Suppose  $\Theta = \bigwedge_{i=0}^q \bigwedge_{j=1}^{s_i} \alpha_{h_{ij}}(a_{ij})$  and for some  $j, k \in \{1, \dots, n\}$ , we have that  $h_{ij} = h_{ik}$ , so  $j E_i k$  for some  $i$ . Since  $\alpha_{h_{ij}} = \alpha_{h_{ik}}$  implies  $\sigma\alpha_{h_{ij}} = \sigma\alpha_{h_{ik}}$ ,  $\sigma$  preserves the  $E_i$ -equivalence of  $j$  and  $k$ . A similar argument gives that if  $j, k \in \{1, \dots, s_i\}$  are not  $E_i$ -equivalent, then  $\sigma\alpha_{h_{ij}} \neq \sigma\alpha_{h_{ik}}$  must hold. Thus we must have that the  $P$ -spectrum of  $\Theta$  is the same as the  $P$ -spectrum of  $\sigma\Theta$ . By SPx for  $w$ ,  $w(\Theta) = w(\sigma\Theta)$ .  $\dashv$

The following counter-example will show that Px does not imply SPx.

**Example 6 (Px does not imply SPx).** Suppose  $\Theta(a_1, \dots, a_5)$ ,  $\Phi(a_1, \dots, a_5)$  are state descriptions of  $L_3$  given by

$$\begin{aligned}\Theta &= (P_1 \wedge P_2 \wedge \neg P_3) \wedge (\neg P_1 \wedge P_2 \wedge P_3) \wedge (\neg P_1 \wedge \neg P_2 \wedge P_3) \\ &\quad \wedge (\neg P_1 \wedge \neg P_2 \wedge \neg P_3) \wedge (P_1 \wedge \neg P_2 \wedge \neg P_3), \\ \Phi &= (P_1 \wedge \neg P_2 \wedge P_3) \wedge (\neg P_1 \wedge P_2 \wedge P_3) \wedge (\neg P_1 \wedge P_2 \wedge \neg P_3) \\ &\quad \wedge (\neg P_1 \wedge P_2 \wedge \neg P_3) \wedge (\neg P_1 \wedge \neg P_2 \wedge P_3).\end{aligned}$$

With the eight atoms of  $L_3$  enumerated in lexicographic order, we obtain

$$\begin{aligned}\Theta &= \alpha_2 \alpha_4 \alpha_5 \alpha_7^2, \\ \Phi &= \alpha_3 \alpha_5 \alpha_6^2 \alpha_7.\end{aligned}$$

Note that there are six permutations of the atoms of  $L_3$  that are induced by Px, and that  $\alpha_2, \alpha_3, \alpha_5$  each have one negated predicate, while  $\alpha_4, \alpha_6, \alpha_7$  each have two negated predicates. Suppose  $\vec{b} \in \mathbb{D}_8$  and let  $\Sigma$  denote the set of all permutations of predicates. Then one can verify easily that  $w$  given by

$$w = \frac{1}{6} \sum_{\sigma \in \Sigma} w_{\sigma \vec{b}}$$

satisfies Px. Notice that  $\Phi$  and  $\Theta$  have the same  $P$ -spectrum. We obtain

$$\begin{aligned}w(\Theta) &= \frac{1}{6} (b_2 b_4 b_5 b_7^2 + b_3 b_4 b_5 b_6^2 + b_2 b_6 b_3 b_7^2 + b_5 b_6 b_3 b_4^2 + b_5 b_7 b_2 b_6^2 + b_3 b_7 b_2 b_6^2), \\ w(\Phi) &= \frac{1}{6} (b_3 b_5 b_7 b_6^2 + b_2 b_5 b_6 b_7^2 + b_5 b_3 b_7 b_4^2 + b_2 b_3 b_4 b_7^2 + b_3 b_2 b_6 b_4^2 + b_5 b_2 b_6 b_4^2).\end{aligned}$$

Now letting

$$\vec{b} = \left\langle \frac{1}{19}, \frac{2}{19}, \frac{4}{19}, \frac{5}{19}, \frac{2}{19}, \frac{3}{19}, \frac{1}{19}, \frac{1}{19} \right\rangle,$$

we clearly have  $\vec{b} \in \mathbb{D}_8$  and we obtain

$$\begin{aligned} w(\Theta) &= \frac{1094}{6 \cdot 19^5}, \\ w(\Phi) &= \frac{1224}{6 \cdot 19^5}, \end{aligned}$$

which clearly gives  $w(\Theta) \neq w(\Phi)$  and thus  $w$  cannot satisfy SPx.

Similarly we can observe that any function satisfying Ax must also satisfy SPx, as we can easily see that any permutation  $\sigma$  of atoms such that  $\gamma_q(\alpha) = \gamma_q(\sigma\alpha)$  preserves  $P$ -spectra. Conversely whenever  $\Theta$  and  $\Phi$  are state descriptions with the same  $P$ -spectrum then there is a bijection between the atoms of  $\Theta$  and those of  $\Phi$  that can be extended to a permutation of atoms. It is easy to see that there are functions satisfying SPx, but not Ax as any permutation that permutes atoms with different numbers of negations can be used to construct a counter-example in just the same fashion as the above example.

### 3 A representation theorem for ULi with SPx

The first step towards the desired representation theorem is defining the class of functions  $v^{\vec{p}, \vec{\tau}}$  that we will show to be the basic building blocks for functions satisfying ULi with SPx. It is apparent from the definition that these functions are generalizations of the  $u^{\vec{p}}$  functions that form the building blocks for Ax + ULi, see e.g. [6]. As was discussed in [4] the classes of  $v^{\vec{p}, \vec{\tau}}$  and  $u^{\vec{p}}$  share a number of properties, of which we will quote (without proof) the most relevant for our purposes.

Let  $\vec{p}, \vec{\tau}$  be countable sequences  $\vec{p} = \langle p_0, p_1, p_2, \dots \rangle$ ,  $\vec{\tau} = \langle \tau_0, \tau_1, \tau_2, \dots \rangle$  and let  $\mathcal{B}$  be the set

$$\mathcal{B} = \left\{ \langle \vec{p}, \vec{\tau} \rangle \mid p_i \geq 0 \text{ for } i \geq 0, p_1 \geq p_2 \geq p_3 \geq \dots \geq p_n \geq \dots, \sum_{i \geq 0} p_i = 1, \right. \\ \left. \tau_j \in [0, 1] \text{ for } j \geq 1 \text{ and } \tau_0 \text{ a normalized } \sigma\text{-additive measure on } [0, 1] \right\}.$$

**Definition 7:**

Let  $L = L_q$  and let  $\langle \vec{p}, \vec{\tau} \rangle \in \mathcal{B}$  and let  $\Theta(\vec{a}) = \Theta(a_1, \dots, a_m)$  be a state description of  $L$  with

$$\Theta(a_1, \dots, a_m) = \bigwedge_{i=1}^m \alpha_{h_i}(a_i).$$

Define the functions  $j_L^{\bar{p}, \bar{\tau}}, v_L^{\bar{p}, \bar{\tau}}$  as follows: Let  $\vec{c}$  be a sequence in  $\mathbb{N}$ . Then

$$j_L^{\bar{p}, \bar{\tau}}(\Theta(\vec{a}), \langle c_1, \dots, c_m \rangle) = \begin{cases} j_L^{\bar{p}, \bar{\tau}}(\Theta^-, \langle c_1, \dots, c_{m-1} \rangle) \cdot p_{c_m} \cdot b_{h_m, c_m} & \text{if } c_m = 0 \text{ or } c_l \neq c_m \text{ for all } l < m, \\ j_L^{\bar{p}, \bar{\tau}}(\Theta^-, \langle c_1, \dots, c_{m-1} \rangle) \cdot p_{c_m} & \text{if } c_l = c_m \neq 0 \text{ for some } l < m \text{ and } h_l = h_m, \\ 0 & \text{otherwise,} \end{cases}$$

with  $\Theta^- = \Theta^-(a_1, \dots, a_{m-1})$  the unique state description such that

$$\Theta(a_1, \dots, a_m) \models \Theta^-(a_1, \dots, a_{m-1})$$

and

$$b_{h_m, c_m} = \begin{cases} \tau_{c_m}^{\gamma_q(h_m)} (1 - \tau_{c_m})^{\gamma_q(h_m)} & \text{if } c_m > 0, \\ \int_{[0,1]} x^{\gamma_q(h_m)} (1-x)^{q-\gamma_q(h_m)} d\tau_0(x) & \text{if } c_m = 0. \end{cases}$$

Define  $v_L^{\bar{p}, \bar{\tau}}$  on state descriptions of  $L$  by

$$v_L^{\bar{p}, \bar{\tau}}(\Theta(\vec{a})) = \sum_{\vec{c}} j_L^{\bar{p}, \bar{\tau}}(\Theta(\vec{a}), \vec{c}). \quad (1)$$

Note that this definition is slightly different than the original definition given in [4]: in the original definition we allowed even the  $\tau_i$  for  $i > 0$  to be normalized  $\sigma$ -additive measures instead of single point measures.

The following theorem lists some properties of the  $v^{\bar{p}, \bar{\tau}}$  functions. The proof is rather lengthy and technical, and in the case of Weak Irrelevance even requires a detour via a different definition for the  $v^{\bar{p}, \bar{\tau}}$ . As we believe the techniques required for the proof do not provide any benefit for the task at hand, and for the sake of brevity, we would like to refer the reader to Chapter 3 of [4], where detailed proofs for the claims are given.

**Theorem 8.** *Let  $\langle \bar{p}, \bar{\tau} \rangle \in \mathcal{B}$ . Then  $v^{\bar{p}, \bar{\tau}}$  is a probability function satisfying the principles  $Ex$ ,  $SPx$ ,  $ULi$  and  $WIP$ .*  $\dashv$

We can now state the representation theorem for functions satisfying  $ULi$  with  $SPx$ . In the proof we will be using methods from non-standard analysis, and in particular Loeb measure theory, to obtain the desired results. See e.g. [2] for details.

**Theorem 9.** *Let  $w$  be a probability function on  $L_q$ . Then  $w$  satisfies  $ULi$  with  $SPx$  if and only if there exists a normalised  $\sigma$ -additive measure  $\mu$  on  $\mathcal{B}$  such that*

$$w = \int_{\mathcal{B}} v^{\bar{p}, \bar{\tau}} d\mu(\langle \bar{p}, \bar{\tau} \rangle). \quad (2)$$

**Proof:** By Theorem 8, it is straightforward to show that any function of the form (2) satisfies ULi with SPx.

For the reverse, let  $w$  be a probability function on  $L_q$  satisfying ULi with SPx. Working in a non-standard universe, let  $\nu \in {}^*\mathbb{N} \setminus \mathbb{N}$ . Then by ULi there exists an extension  $w^{L_\nu}$  of  $w$  to  $L_\nu$ .

We can write  $w^{L_\nu}$  as

$$w^{L_\nu}(\Theta(a_1, \dots, a_n)) = \sum_{\substack{\Phi(a_1, \dots, a_n) \in \text{SDL}_\nu \\ \Phi \models \Theta}} w^{L_\nu}(\Phi(a_1, \dots, a_n)). \quad (3)$$

Since  $w^{L_\nu}$  satisfies SPx, we can rephrase (3): fix some  $P$ -spectrum  $\hat{\nu}$  of a state description  $\Upsilon(a_1, \dots, a_\nu)$  of  $L_\nu$ . Then for each state description  $\Phi(a_1, \dots, a_\nu)$  such that  $\text{PSpec}(\Phi) = \text{PSpec}(\Upsilon) = \hat{\nu}$  we have  $w^{L_\nu}(\Phi) = w^{L_\nu}(\Upsilon)$ .

Letting  $\overline{\Upsilon} = \{\Phi(a_1, \dots, a_\nu) \mid \text{PSpec}(\Phi) = \text{PSpec}(\Upsilon)\}$  for a fixed state description  $\Upsilon$ , and partitioning the set of all state descriptions of  $L_\nu$  for constants  $a_1, \dots, a_\nu$  by a choice of  $\overline{\Upsilon}$  we obtain

$$\begin{aligned} w^{L_\nu}(\Theta(a_1, \dots, a_n)) &= \sum_{\overline{\Upsilon}} \sum_{\substack{\Phi \in \overline{\Upsilon} \\ \Phi \models \Theta}} w^{L_\nu}(\Phi(a_1, \dots, a_n)) \\ &= \sum_{\overline{\Upsilon}} |\{\Phi \in \overline{\Upsilon} \mid \Phi \models \Theta\}| \cdot w^{L_\nu}(\Upsilon(a_1, \dots, a_n)) \\ &= \sum_{\overline{\Upsilon}} \frac{|\{\Phi \in \overline{\Upsilon} \mid \Phi \models \Theta\}|}{|\overline{\Upsilon}|} \cdot |\overline{\Upsilon}| \cdot w^{L_\nu}(\Upsilon) \\ &= \sum_{\overline{\Upsilon}} \frac{|\{\Phi \in \overline{\Upsilon} \mid \Phi \models \Theta\}|}{|\overline{\Upsilon}|} \cdot w^{L_\nu}(\bigvee \overline{\Upsilon}), \end{aligned}$$

where  $\bigvee \overline{\Upsilon}$  is the disjunction over all state descriptions in  $\overline{\Upsilon}$ .

We will show that  $\frac{|\{\Phi \in \overline{\Upsilon} \mid \Phi \models \Theta\}|}{|\overline{\Upsilon}|}$  is a probability function of the form  $v^{\overline{p}, \overline{\tau}}$  for some  $\langle \overline{p}, \overline{\tau} \rangle \in \mathcal{B}$ .

First note that if  $\Phi \in \overline{\Upsilon}$  then there exists a permutation  $\sigma$  of the atoms of  $L_\nu$  preserving  $P$ -spectra and a permutation  $\rho$  of the constants  $a_1, \dots, a_\nu$  such that  $\Phi(a_1, \dots, a_\nu) = \sigma\Upsilon(a_{\rho(1)}, \dots, a_{\rho(\nu)})$ . We can then view  $\frac{|\{\Phi \in \overline{\Upsilon} \mid \Phi \models \Theta\}|}{|\overline{\Upsilon}|}$  as the probability of obtaining some  $\Phi(a_1, \dots, a_\nu) \models \Theta(a_1, \dots, a_n)$  with  $\Phi \in \overline{\Upsilon}$  by randomly picking permutations  $\sigma$  and  $\rho$  such that  $\Phi(a_1, \dots, a_\nu) = \sigma\Upsilon(a_{\rho(1)}, \dots, a_{\rho(\nu)})$ .

Fix some  $\Upsilon \in \text{SDL}_\nu$  with  $\text{PSpec}(\Upsilon) = \hat{\nu} = \langle s_0, s_{1,1}, \dots, s_{1,i_1}, s_{2,1}, \dots, s_{2,i_2}, \dots, s_\nu \rangle$ , where



the  $s_{i,j}$  are the sizes of the equivalence classes in the  $P$ -spectrum of  $\Upsilon$ .<sup>4</sup> Then the probability of picking a constant of a particular equivalence class at random is  $\frac{s_{i,j}}{\nu}$ , where  $s_{i,j}$  is the size of this equivalence class according to  $\text{PSpec}(\Upsilon)$ . Let  $p_1 \geq p_2 \geq p_3 \geq \dots \geq p_N$  be the list of these probabilities, in decreasing order, and let  $s'_1, \dots, s'_N$  be an enumeration of the equivalence classes such that  $p_i = \frac{s'_i}{\nu}$ . Here,  $N$  is the number of equivalence classes in  $\hat{\nu}$ . Note that  $N$  may very well be infinite. Then the probability of randomly picking (with replacement) constants  $a_1, \dots, a_n$  from equivalence classes  $s'_{c_1}, \dots, s'_{c_n}$ , for  $1 \leq c_1, \dots, c_n \leq N$ , is  $\prod_{i=1}^n p_{c_i}$ .<sup>5</sup>

We now need to take care of the specific atom picked for each of the classes  $s'_{c_1}, \dots, s'_{c_n}$ . Again, as we are going to take standard parts and thus the difference between picking atoms with and without replacement will disappear, we can treat picking an atom to represent each class as independent of each other. Note that the only restriction for picking an atom is that if the original atom in  $\Upsilon$  representing this class has  $j$  negated predicates occurring, so must the atom we pick at random have  $j$  predicates in order to preserve the  $P$ -spectrum of  $\Upsilon$ . The atom we pick does not even have to occur in  $\Upsilon$ . We can view randomly picking an atom in this way just as picking predicates without replacement from the original atom, determining the atom we want to replace it with in this way.

Consider the equivalence class  $s'_{c_m}$  for some  $m \leq n$ . Suppose that the atom of  $L_\nu$  this equivalence class represents has  $j$  negated predicates occurring. Suppose that  $\Theta(a_1, \dots, a_n) \models \alpha_{h_m}(a_m)$  with  $\gamma_q(\alpha_{h_m}) = r$ . Then the probability that picking an atom  $\beta$  of  $L_\nu$  with  $\gamma_\nu(\beta) = j$  and  $\beta(x) \models \alpha_{h_m}(x)$  is determined as follows: picking predicates randomly from  $\beta$  without replacement and if the predicate  $P_i$  for  $1 \leq i \leq q$  occurs positively in  $\alpha_{h_m}$ , we obtain a factor  $\frac{\nu-j-k_i}{\nu-(i-1)}$ , where  $k_i$  is the number of predicates  $P_t$ ,  $1 \leq t < i$  that occur positively in  $\alpha_{h_m}$ . Similarly if  $P_i$  occurs negatively we obtain a factor  $\frac{j-k_i}{\nu-(i-1)}$  with  $k_i$  the number of predicates occurring negatively in  $\alpha_{h_m}$ . Once we have picked  $q$  predicates this way, we are left with an arbitrary choice of predicates to pick, resulting in a factor of 1 for the choice of  $P_t$ ,  $t > q$ . By commutativity, we obtain the probability as

$$\prod_{i=0}^{r-1} \frac{j-i}{\nu-i} \cdot \prod_{i=0}^{q-r-1} \frac{\nu-j-i}{\nu-i-r}. \quad (4)$$

As we will be taking standard parts and  $i$  only takes finite values, we may write (4) as

$$\left(\frac{j}{\nu}\right)^r \cdot \left(\frac{\nu-j}{\nu}\right)^{q-r} = \tau^r (1-\tau)^{q-r}, \quad (5)$$

---

<sup>4</sup>Note that  $s_0$  and  $s_\nu$  only have a single index, as there can be only one equivalence class of constants instantiating an atom with no negations or only negations occurring, respectively.

<sup>5</sup>Just as in the proof of the representation theorem for ULi (see [5]) it will suffice to consider picking constants with replacement, as the difference to picking constants without replacement will disappear once we take standard parts.

for some  $\tau \in {}^*[0, 1]$ , as the difference between (4) and (5) will disappear once we have taken standard parts.

Thus for the probability that a fixed  $\Phi \in \overline{\Upsilon}$  is such that  $\Phi \models \bigwedge_{i=1}^n \alpha_{h_i}(a_i)$  we obtain a factor of  $\prod_{i=1}^n p_{c_i}$ , where  $s'_{c_i}$  is the equivalence class of  $a_i$  in the  $P$ -spectrum of  $\Phi$ . Furthermore we obtain a factor of  $\tau_k^{\gamma_q(\alpha_{h_i})} (1 - \tau_k)^{q - \gamma_q(\alpha_{h_i})}$  for each  $k \in \{c_1, \dots, c_n\}$ . The reason we only have one occurrence of this factor is that once we have picked an atom for one constant instantiating a particular equivalence class of the  $P$ -spectrum, each further constant from the same class must instantiate the same atom. For the same reason, the probability must be 0 if we have  $c_i = c_j$  for some  $1 \leq i, j \leq n$  but  $\alpha_{h_i}$  and  $\alpha_{h_j}$  are two different atoms. In contrast to the Px case we can treat the picking of predicates described above as independent for each equivalence class of atoms: we are in fact just picking atoms of  $L_\nu$  with the correct number of negations occurring while ensuring their restriction to  $L_q$  yields the desired atom. As the probability of picking the same atom of  $L_\nu$  for any two equivalence classes is infinitesimal, we can waive the difference between picking these atoms with and without replacement, resulting in the probability of each atom being picked independent of the choice of equivalence classes. Thus we obtain

$$\prod_{i=1}^n p_{c_i} \cdot \prod_{k \in \{c_1, \dots, c_n\}} \tau_k^{\gamma_q(\alpha_{h_i})} (1 - \tau_k)^{q - \gamma_q(\alpha_{h_i})}$$

as the probability, in case that whenever  $c_i = c_j$  then  $h_i = h_j$ . But this is just a non-standard version of  $j^{\overline{p}, \overline{\tau}}(\bigwedge_{i=1}^n \alpha_{h_i}(a_i), \langle c_1, \dots, c_n \rangle)$  for  $\overline{p} = \langle 0, p_1, p_2, \dots \rangle$  and  $\overline{\tau} = \langle \tau_0, \tau_1, \tau_2, \dots \rangle$  for an arbitrary measure  $\tau_0$ .

We now obtain

$$\frac{|\{\Phi \in \overline{\Upsilon} \mid \Phi \models \Theta\}|}{|\overline{\Upsilon}|} = \sum_{\substack{\langle c_1, \dots, c_n \rangle \\ c_i \geq 1}} {}^*j^{\overline{p}, \overline{\tau}}(\Theta(a_1, \dots, a_n), \langle c_1, \dots, c_n \rangle) = {}^*v^{\overline{p}, \overline{\tau}}(\Theta(a_1, \dots, a_n)),$$

for a fixed  $\Theta$  such that  $\text{PSpec}(\Theta) = \hat{n}$  and the  $*$  indicating the non-standard versions of the functions defined in Definition 7.

Define a measure  $\mu$  on the  $P$ -spectra of state descriptions of  $L_\nu$  by  $\mu(\{\hat{\nu}\}) = w^{L_\nu}(\bigvee \overline{\Upsilon})$  and we obtain

$$\begin{aligned} \sum_{\hat{\nu}} \frac{|\{\Phi \in \overline{\Upsilon} \mid \Phi \models \Theta\}|}{|\overline{\Upsilon}|} \cdot w^{L_\nu}(\bigvee \overline{\Upsilon}) &= \int_{\hat{\nu}} \frac{|\{\Phi \in \overline{\Upsilon} \mid \Phi \models \Theta\}|}{|\overline{\Upsilon}|} d\mu(\hat{\nu}) \\ &= \int_{\hat{\nu}} {}^*v^{\overline{p}, \overline{\tau}}(\hat{n}) d\mu(\hat{\nu}). \end{aligned}$$

Now taking standard parts we obtain

$$\begin{aligned} w^{L_q}(\Theta(a_1, \dots, a_n)) &= \left( \int_{\hat{\nu}}^{\circ} {}^*v^{\bar{p}, \bar{\tau}}(\Theta(a_1, \dots, a_n)) d\mu(\hat{\nu}) \right) \\ &= \left( \int_{\langle \bar{p}, \bar{\tau} \rangle}^{\circ} {}^*v^{\bar{p}, \bar{\tau}}(\Theta(a_1, \dots, a_n)) d\mu(\langle \bar{p}, \bar{\tau} \rangle) \right), \end{aligned} \quad (6)$$

as  $\langle \bar{p}, \bar{\tau} \rangle$  represents  $\hat{\nu}$ .

We can now use Loeb measures to move the operation of taking standard parts inside the integral and claim that  ${}^{\circ}({}^*v^{\bar{p}, \bar{\tau}}(\Theta(a_1, \dots, a_n))) = v^{\bar{p}', \bar{\tau}'}(\Theta(a_1, \dots, a_n))$  for some standard versions  $\bar{p}'$  of  $\bar{p}$  and  $\bar{\tau}'$  for  $\bar{\tau}$ .

It remains to find  $\langle \bar{p}', \bar{\tau}' \rangle \in \mathcal{B}$  such that

$${}^{\circ}({}^*v^{\bar{p}, \bar{\tau}}(\Theta(a_1, \dots, a_n))) = v^{\bar{p}', \bar{\tau}'}(\Theta(a_1, \dots, a_n)) \quad (7)$$

First notice that  ${}^*j^{\bar{p}, \bar{\tau}}(\Theta(a_1, \dots, a_n), \vec{c})$  is a finite product for each  $\vec{c}$  and so we can move the standard part operation all the way to the individual factors. This provides us with an obvious candidate for  $\bar{\tau}'$ : take  $\tau'_i = {}^{\circ}\tau_i$  for each  $i \in \mathbb{N}$ ,  $i \geq 1$ .

For the  $p_i$  there are three cases to consider. If there are only finitely many equivalence classes in  $\hat{\nu}$ , there exists some  $n \in \mathbb{N}$  such that  $p_i = 0$  for each  $i > n$ . In this case we have  $\sum_{i=1}^n {}^{\circ}p_i = 1$  (as  $\sum_{i \in \mathbb{N}} p_i = 1$ ) and we can take  $p'_i = {}^{\circ}p_i$  for each  $1 \leq i \leq n$  and let  $p'_i = 0$  otherwise. Then clearly we have

$${}^{\circ}({}^*j^{\bar{p}, \bar{\tau}}(\Theta(a_1, \dots, a_n), \vec{c})) = j^{\bar{p}', \bar{\tau}'}(\Theta(a_1, \dots, a_n), \vec{c}) \quad (8)$$

and (7) holds for  $\langle \bar{p}', \bar{\tau}' \rangle$  as there are only finitely many instances of  $j^{\bar{p}', \bar{\tau}'}$  occurring.

Otherwise we have that  $p_i > 0$  for each  $i \in \mathbb{N}$ . Suppose that in this case we have  $\sum_{i \in \mathbb{N}} {}^{\circ}p_i = 1$ , i.e. there is no weight carried by the  $p_i$  with non-standard indices  $i \in {}^*\mathbb{N} \setminus \mathbb{N}$ . Then we can take  $p'_i = {}^{\circ}p_i$  for each  $i \in \mathbb{N}$ ,  $i \geq 1$  and  $p_0 = 0$  for (8) to hold: we can interpret  $v^{\bar{p}, \bar{\tau}}(\Theta(a_1, \dots, a_n), \vec{c})$  as an instance of integration by a discrete measure: Let  $\rho$  be the product measure giving weight  $\prod_{i=1}^n p_{c_i}$  to the point  $\vec{c} \in {}^*\mathbb{N}^n$  and let  $f$  be the function defined by

$$f(\Theta(a_1, \dots, a_n), \vec{c}) = \prod_{i \in \{c_1, \dots, c_n\}} \tau_i^{\gamma_q(h_i)} (1 - \tau_i)^{q - \gamma_q(h_i)}.$$

Then we can write

$${}^*v^{\bar{p}, \bar{\tau}}(\Theta(a_1, \dots, a_n), \vec{c}) = \sum_{\vec{c} \in {}^*\mathbb{N}^n} j^{\bar{p}, \bar{\tau}}(\Theta(a_1, \dots, a_n), \vec{c}) = \int_{{}^*\mathbb{N}^n} f(\Theta(a_1, \dots, a_n), \vec{c}) d\rho(\vec{c}). \quad (9)$$

Now taking the standard part of the integral in (9) and applying Loeb measure theory in this situation, we obtain as Loeb measure  $\rho^L$  the measure on  $\mathbb{N}^n$  giving weight  $\prod_{i=1}^n p'_{c_i}$

to the point  $\vec{c} \in \mathbb{N}^n$ , observing that whenever  $\vec{c} \in {}^*\mathbb{N}^n$  is such that for some  $j \leq n$ ,  $c_j \in {}^*\mathbb{N} \setminus \mathbb{N}$ , we have  ${}^\circ(\prod_{i=1}^n p_{c_i}) = 0$  and thus (7) holds.

Lastly suppose that  $\sum_{i \in \mathbb{N}} {}^\circ p_i < 1$ . In this case we immediately have that while  ${}^*v^{\vec{p}, \vec{\tau}}(\top) = 1$ , simply taking standard parts of each  $p_i$  as above will lead to  $v^{\vec{p}', \vec{\tau}'}(\top) < 1$ . The obvious problem now is to attribute the weight ‘lost’ by simply taking standard parts to the  $p'_i$  with standard indices. As  $p_0$  has not been used yet, we will put all this weight into  $p'_0$  and it remains to find a measure  $\tau_0$  such that the equation (7) holds.

In more detail, let  $p'_i = {}^\circ p_i$  for each  $i \in \mathbb{N}$ ,  $i \geq 1$  and let  $p'_0 = 1 - \sum_{i \in \mathbb{N}} {}^\circ p_i$ . Fix some  $\vec{c}$  such that there exists  $j$  with  ${}^\circ p_{c_j} = 0$  and  $p_{c_j} > 0$ . Let  $\mathcal{C}(\vec{c}, j)$  be the collection

$$\mathcal{C}(\vec{c}, j) = \{\vec{e}' \mid \forall i (i \neq j \rightarrow e_i = c_i) \wedge (i = j \rightarrow p_{e_i} > 0 \wedge {}^\circ p_{e_i} = 0)\},$$

i.e. the sequences in  $\mathcal{C}(\vec{c}, j)$  only differ in the  $j$ th component, and each of the disappearing  $p_i$  occurs in one of the sequences.

Then we obtain

$$\begin{aligned} & {}^\circ \left( \sum_{\vec{e}' \in \mathcal{C}(\vec{c}, j)} \prod_{i=1}^n p_{e_i} \cdot \prod_{k \in \{e_1, \dots, e_n\}} \tau_k^{\gamma_q(\hat{h}_k)} (1 - \tau_k)^{q - \gamma_q(\hat{h}_k)} \right) = \\ & {}^\circ \left( \prod_{i=1}^n p_{c_i} \cdot \prod_{k \in \{c_1, \dots, c_n\} \setminus \{c_j\}} \tau_k^{\gamma_q(\hat{h}_k)} (1 - \tau_k)^{q - \gamma_q(\hat{h}_k)} \right) \cdot {}^\circ \left( \sum_{\substack{p_j > 0 \\ {}^\circ p_j = 0}} p_j \cdot \tau_j^{\gamma_q(h_j)} (1 - \tau_j)^{q - \gamma_q(h_j)} \right), \end{aligned}$$

where  $\hat{h}_k$  here is that  $h_s$  such that  $k = c_s$  ( $k = e_s$ , respectively).

Let  $\tau_0$  be a measure on  ${}^*[0, 1]$  such that  $\tau_0$  gives weight  $p_j/p'_0$  to the point  $\tau_j$  for each  $j$  such that  $p_j > 0$  and  ${}^\circ p_j = 0$ , and weight 0 to all other points. Then we obtain

$$\begin{aligned} & {}^\circ \left( \sum_{\substack{p_j > 0 \\ {}^\circ p_j = 0}} p_j \cdot \tau_j^{\gamma_q(h_j)} (1 - \tau_j)^{q - \gamma_q(h_j)} \right) = p'_0 \cdot {}^\circ \left( \int_{{}^*[0, 1]} x^{\gamma_q(h_j)} (1 - x)^{q - \gamma_q(h_j)} d\tau_0(x) \right) \\ & = p'_0 \int_{[0, 1]} x^{\gamma_q(h_j)} (1 - x)^{q - \gamma_q(h_j)} d\tau_0^L(x), \end{aligned} \quad (10)$$

where  $\tau_0^L$  is the corresponding Loeb measure to  $\tau_0$ . We can continue in this way until all such  $p_i$  have been collected into a factor of the form (10).

Letting  $\tau'_0 = \tau_0^L$  a straightforward calculation shows that

$${}^\circ {}^*v^{\vec{p}, \vec{\tau}}(\Theta) = v^{\vec{p}', \vec{\tau}'}(\Theta), \quad (11)$$

as we can now reduce the left-hand side to the situation given in (9), with  $p_0 = p'_0$ ,  $\tau_0$  as defined above, and  $\sum_{i \in \mathbb{N}} {}^\circ p_i = 1$  : as the  $p_i > 0$  with  ${}^\circ p_i = 0$  are now collected into  $p_0$ , just let these  $p_i = 0$ .

Going back to (6), we obtain together with (11)

$$\begin{aligned} w^{L_q}(\Theta(a_1, \dots, a_n)) &= {}^\circ \left( \int_{\langle \bar{p}, \bar{\tau} \rangle} {}^* v^{\bar{p}, \bar{\tau}}(\Theta(a_1, \dots, a_n)) d\mu(\langle \bar{p}, \bar{\tau} \rangle) \right) \\ &= \int_{\langle \bar{p}', \bar{\tau}' \rangle} v^{\bar{p}', \bar{\tau}'}(\Theta(a_1, \dots, a_n)) d\mu^L(\langle \bar{p}', \bar{\tau}' \rangle), \end{aligned}$$

where  $\mu^L$  is the Loeb measure corresponding to  $\mu$ . ⊣

**Lemma 10.** *The  $v^{\bar{p}, \bar{\tau}}$  are the only functions satisfying ULi with SPx + WIP.*

**Proof:** Let  $w$  be a probability function satisfying ULi with SPx + WIP and let  $\vartheta \in QFSL$ . Using ULI, extend  $w$  to some  $w'$  on a language  $L'$  large enough so that we can permute the predicates and constants of  $\vartheta$  to obtain  $\vartheta' \in QFSL'$  with no predicates or constants in common with  $\vartheta$ . As SPx implies Px, and by ULi we obtain  $w(\vartheta) = w'(\vartheta) = w'(\vartheta')$ . Then we obtain

$$\begin{aligned} 0 &= 2(w'(\vartheta \wedge \vartheta') - w'(\vartheta) \cdot w'(\vartheta')) \\ &= \int_{\mathcal{B}} v_{L'}^{\bar{\tau}, \bar{\rho}}(\vartheta \wedge \vartheta') d\mu(\langle \bar{\tau}, \bar{\rho} \rangle) - 2 \int_{\mathcal{B}} v_{L'}^{\bar{\tau}, \bar{\rho}}(\vartheta) d\mu(\langle \bar{\tau}, \bar{\rho} \rangle) \cdot \int_{\mathcal{B}} v_{L'}^{\bar{p}, \bar{\tau}}(\vartheta') d\mu(\langle \bar{p}, \bar{\tau} \rangle) \\ &\quad + \int_{\mathcal{B}} v_{L'}^{\bar{p}, \bar{\tau}}(\vartheta \wedge \vartheta') d\mu(\langle \bar{p}, \bar{\tau} \rangle) \\ &= \int_{\mathcal{B}} \int_{\mathcal{B}} (v_{L'}^{\bar{\tau}, \bar{\rho}}(\vartheta)^2 - 2v_{L'}^{\bar{\tau}, \bar{\rho}}(\vartheta) \cdot v_{L'}^{\bar{p}, \bar{\tau}}(\vartheta) + v_{L'}^{\bar{p}, \bar{\tau}}(\vartheta)^2) d\mu(\langle \bar{\tau}, \bar{\rho} \rangle) d\mu(\langle \bar{p}, \bar{\tau} \rangle) \\ &= \int_{\mathcal{B}} \int_{\mathcal{B}} (v_{L'}^{\bar{\tau}, \bar{\rho}}(\vartheta) - v_{L'}^{\bar{p}, \bar{\tau}}(\vartheta))^2 d\mu(\langle \bar{\tau}, \bar{\rho} \rangle) d\mu(\langle \bar{p}, \bar{\tau} \rangle), \end{aligned}$$

by the Representation Theorem 9. As the function under the integral is non-negative we must have a measure 1 subset of  $\mathcal{B}$  such that  $v_{L'}^{\bar{p}, \bar{\tau}}$  is constant on this set for each  $\vartheta \in QFSL$  and therefore we must have  $w' = v_{L'}^{\bar{p}, \bar{\tau}}$  for any  $\langle \bar{p}, \bar{\tau} \rangle$  in this set. By ULi we obtain that  $w = w' \upharpoonright SL = v_L^{\bar{p}, \bar{\tau}}$ , as required. ⊣

## 4 The General Representation Theorem

Just as for the probability functions satisfying Predicate and Atom Exchangeability<sup>6</sup>, respectively, we obtain a similar result for Strong Predicate Exchangeability:

---

<sup>6</sup>See e.g. [5] for the result concerning Predicate Exchangeability and [6, Chapter 34] for the result concerning Atom Exchangeability.

**Theorem 11.** *Let  $w$  be a probability function on  $L$ . Then  $w$  satisfies SPx if and only if there are  $\lambda \geq 0$  and probability functions  $w_1, w_2$  satisfying ULi with SPx such that*

$$w = (1 + \lambda)w_1 - \lambda w_2.$$

In the proof, we will use a slightly different formulation<sup>7</sup> for the  $v_{n,L_q}^{\bar{p},\bar{\tau}}$  functions: Let  $v_{n,L_q}^{\bar{p},\bar{\tau}}$  be defined by

$$v_{n,L_q}^{\bar{p},\bar{\tau}} = \sum_{e \in Z_n^q} \prod_{r=1}^n \tau_r^{\gamma_q(e(r))} (1 - \tau_r)^{q - \gamma_q(e(r))} \cdot w_{e(\bar{p})}, \quad (12)$$

where  $Z_n^q$  is the set of all functions  $e : \{1, \dots, n\} \rightarrow \{1, \dots, 2^q\}$  and  $e(\bar{p}) \in \mathbb{D}_{2^q}$  is given by

$$e(\bar{p}) = \langle f^q(1, 0)R_{\bar{p},n} + \sum_{e(i)=1} p_i, f^q(2, 0)R_{\bar{p},n} + \sum_{e(i)=2} p_i, \dots, f^q(2^q, 0)R_{\bar{p},n} + \sum_{e(i)=2^q} p_i \rangle,$$

with  $R_{\bar{p},n} = 1 - \sum_{i=1}^n p_i$  and  $f^q(s, 0) = \int_{[0,1]} x^{\gamma_q(s)} (1 - x)^{q - \gamma_q(s)} d\tau_0(x)$ . It is easy to verify that if  $p_i = 0$  for  $i > n$ , then  $v_{n,L_q}^{\bar{p},\bar{\tau}} = v_{L_q}^{\bar{p},\bar{\tau}}$ . As usual, we will omit  $L_q$  whenever it is understood from the context.

In the proof we will replicate the methods used for the analogous theorem for probability functions satisfying Predicate Exchangeability, see e.g. Lemma 13 and Theorem 14 in [5].

**Proof:** Suppose that  $w = (1 + \lambda)w_1 - \lambda w_2$  for some  $\lambda \geq 0$  and  $w_1, w_2$  probability functions satisfying ULi with SPx. Then it is straightforward to check that  $w$  is a probability function satisfying SPx, but not necessarily ULi.

So suppose that  $w$  is a probability function satisfying SPx. Let  $\vec{x} \in \mathbb{D}_{2^q}$  and consider the function  $w_{\vec{x}}$ . We can close this function under SPx by letting

$$z_{\vec{x}} = \frac{1}{|\Sigma|} \sum_{\sigma \in \Sigma} w_{\sigma(\vec{x})}, \quad (13)$$

where  $\Sigma$  is the set of all permutations of  $\{1, \dots, 2^q\}$  that preserve  $P$ -spectra.

It is then easy to check that we obtain a de Finetti representation of  $w$  of the form

$$w = \int_{\mathbb{D}_{2^q}} z_{\vec{x}} d\mu(\vec{x}). \quad (14)$$

So suppose that  $\mu$  is a measure putting all weight on a singleton  $\{\vec{x}\}$ , i.e.  $w = z_{\vec{x}}$  for some  $\vec{x} \in \mathbb{D}_{2^q}$ .

---

<sup>7</sup>In fact, the functions  $v_{n,L_q}^{\bar{p},\bar{\tau}}$  are instrumental to the proof of Theorem 8.

Pick  $\bar{p}$  such that  $p_0 = 0$ ,  $p_i = x_i$  for  $i = 1, \dots, 2^q$  and  $p_i = 0$  for  $i > 2^q$ . Then for arbitrary  $\bar{\tau}$ ,  $v_{2^q}^{\bar{p}, \bar{\tau}}$  will have  $z_{\bar{x}}$  occurring in its representation (12), as clearly  $e \in Z_n^q$  such that  $e(i) = i$  will result in  $w_{\bar{x}} = w_{e(\bar{p})}$  occurring, and for each permutation  $\sigma$  such that  $w_{\sigma(\bar{x})}$  occurs in  $z_{\bar{x}}$  the function  $e' = e \circ \sigma \in Z_n^q$  will ensure that  $w_{\sigma(\bar{x})}$  occurs in  $v_{2^q}^{\bar{p}, \bar{\tau}}$  as well.

Consider the factor  $\prod_{r=1}^{2^q} \tau_r^{\gamma_q(e(r))} (1 - \tau_r)^{q - \gamma_q(e(r))}$  accompanying  $w_{e(\bar{p})}$ . If  $\sigma$  is a permutation of  $\{1, \dots, 2^q\}$  that preserves  $P$ -spectra, then for  $e' = e \circ \sigma$  we can easily check that

$$\prod_{r=1}^{2^q} \tau_r^{\gamma_q(e(r))} (1 - \tau_r)^{q - \gamma_q(e(r))} = \prod_{r=1}^{2^q} \tau_r^{\gamma_q(e'(r))} (1 - \tau_r)^{q - \gamma_q(e'(r))}, \quad (15)$$

as the atoms  $\alpha_{e(r)}$ ,  $\alpha_{e'(r)}$  have the same number of negations.

So we have each  $w_{\sigma(\bar{x})}$  occurring in  $z_{\bar{x}}$  occurring in  $v_{2^q}^{\bar{p}, \bar{\tau}}$  with the same factor. Similarly, for each  $\bar{x} \in \mathbb{D}_{2^q}$  such that  $w_{\bar{x}}$  occurs in  $v_{2^q}^{\bar{p}, \bar{\tau}}$ , all the other  $w_{\bar{y}}$  occurring in  $z_{\bar{x}}$  also occur in  $v_{2^q}^{\bar{p}, \bar{\tau}}$ , so we can write

$$v_{2^q}^{\bar{p}, \bar{\tau}} = \sum_{e \in E} a_{e, \bar{\tau}} \cdot S_e \cdot z_{e(\bar{p})}, \quad (16)$$

where  $S_e$  is a normalizing factor,  $a_{e, \bar{\tau}}$  is the factor  $\prod_{r=1}^{2^q} \tau_r^{\gamma_q(e(r))} (1 - \tau_r)^{q - \gamma_q(e(r))}$  and  $E$  is the collection of representatives of equivalence classes of  $Z_n^q$  under the equivalence relation  $\sim_{Z_n^q}$  defined by

$$e \sim_{Z_n^q} e' \Leftrightarrow w_{e(\bar{p})}, w_{e'(\bar{p})} \text{ occur in the same } z_{\bar{x}}.$$

Note that while  $v_{2^q}^{\bar{p}, \bar{\tau}}$  depends on both  $\bar{p}$  and  $\bar{\tau}$ , the functions  $z_{e(\bar{p})}$  only depend on  $\bar{p}$  while  $a_{e, \bar{\tau}}$  only depends on  $\bar{\tau}$ .

Fixing an enumeration of  $E$ , fix some  $\bar{\tau}_e$  for each  $e \in E$  and let  $A$  be the  $E \times E$  matrix with entry  $a_{\langle e, f \rangle} = a_{f, \bar{\tau}_e}$ . We then obtain the equation

$$\begin{pmatrix} \vdots \\ v_{2^q}^{\bar{p}, \bar{\tau}_e} \\ \vdots \end{pmatrix} = A \cdot \begin{pmatrix} \vdots \\ S_e \cdot z_{e(\bar{p})} \\ \vdots \end{pmatrix}. \quad (17)$$

We will show that the entries of  $A$ , i.e. the  $\bar{\tau}_e$ , can be picked such that  $A$  is regular.

Letting  $t = |E|$ , define the matrix  $A_{\langle i_1, \dots, i_j \rangle}$  for  $1 \leq i_1 < i_2 < \dots < i_j \leq t$  the  $j \times j$  submatrix of  $A$  obtained by taking the  $i_1, \dots, i_j$ 'th rows and columns of  $A$ . By induction on  $j$ , we will show that the  $\bar{\tau}_{s, i_t}$  can be picked such that  $A_{\langle i_1, \dots, i_j \rangle}$  is regular. It will suffice to obtain for each row products of the form  $\prod_{r=1}^{2^q} x_r^s y_r^{q-s}$ , not necessarily  $x + y = 1$ , nor

even  $x, y < 1$ , as we can add a factor  $(x_r + y_r)^q$  for each  $r \in \{1, \dots, 2^q\}$  to the entry in  $A$  to obtain the required form.

For  $j = 1$ , this is trivial: just pick each  $\tau_s$  to be neither 0 nor 1. So suppose  $j = k + 1$  for some  $k \geq 1$ . The polynomial  $\prod_{r=1}^{2^q} x_r^{s_r} y_r^{q-s_r}$  takes its maximum value on  $\mathbb{D}_{2^{2q+1}}$  at  $x_r = s_r/2^{q+1}$ ,  $y_r = q - s_r/2^{q+1}$ ,  $r \in \{1, \dots, 2^q\}$ . Considering the previously fixed enumeration of  $E$ , there exists some  $e$  such that

$$\prod_{r=1}^{2^q} (s_{r,e} 2^{-q-1})^{s_{r,e}} ((q - s_{r,e}) 2^{-q-1})^{q-s_{r,e}} > \prod_{r=1}^{2^q} (s_{r,e} 2^{-q-1})^{s_{r,f}} ((q - s_{r,e}) 2^{-q-1})^{q-s_{r,f}}$$

for any  $f \neq e$ . For if not, then for some  $f \neq e$  we have

$$\begin{aligned} \prod_{r=1}^{2^q} (s_{r,e} 2^{-q-1})^{s_{r,e}} ((q - s_{r,e}) 2^{-q-1})^{q-s_{r,e}} &\leq \prod_{r=1}^{2^q} (s_{r,e} 2^{-q-1})^{s_{r,f}} ((q - s_{r,e}) 2^{-q-1})^{q-s_{r,f}} \\ &< \prod_{r=1}^{2^q} (s_{r,f} 2^{-q-1})^{s_{r,f}} ((q - s_{r,f}) 2^{-q-1})^{q-s_{r,f}} \end{aligned}$$

and continuing in this way, we arrive at a contradiction.

By the inductive hypothesis there are choices for the  $x_{r,i_m}, y_{r,i_m}$  for  $i_m \in \{i_1, \dots, i_j\} \setminus \{i_e\}$  such that the sub-matrix  $A_{\langle i_1, \dots, i_{e-1}, i_{e+1}, \dots, i_j \rangle}$  is regular. Treating the  $x_r, y_r$  as unknowns for the moment we obtain for the determinant of  $A_{\langle i_1, \dots, i_j \rangle}$  an expression of the form

$$\begin{aligned} \det(A_{\langle i_1, \dots, i_j \rangle}) &= \\ \pm \prod_{r=1}^{2^q} x_r^{s_{r,e}} y_r^{q-s_{r,e}} \cdot \det(A_{\langle i_1, \dots, i_{e-1}, i_{e+1}, \dots, i_j \rangle}) &+ \sum_{f \in \{i_1, \dots, i_j\} \setminus \{i_e\}} \prod_{r=1}^{2^q} x_r^{s_{r,f}} y_r^{q-s_{r,f}} \cdot (\pm \det(A_f)), \end{aligned} \tag{18}$$

for some choices of  $\pm$ , where the  $A_f$  are the corresponding sub-matrices of  $A_{\langle i_1, \dots, i_j \rangle}$ . Picking  $x_r = (s_{r,e} 2^{-q-1})^g$ ,  $y_r = ((q - s_{r,e}) 2^{-q-1})^g$  for large enough  $g > 0$  the dominant term of (18) becomes

$$\prod_{r=1}^{2^q} x_r^{s_{r,e}} y_r^{q-s_{r,e}} \cdot \det(A_{\langle i_1, \dots, i_{e-1}, i_{e+1}, \dots, i_j \rangle}),$$

and as for this choice of  $x_r, y_r$  we clearly have  $\prod_{r=1}^{2^q} x_r^{s_{r,e}} y_r^{q-s_{r,e}} > 0$  and by the inductive hypothesis  $\det(A_{\langle i_1, \dots, i_{e-1}, i_{e+1}, \dots, i_j \rangle}) \neq 0$ , we get that  $\det(A_{\langle i_1, \dots, i_j \rangle}) \neq 0$ . Now introducing factors  $\prod_{r=1}^{2^q} (x_r + y_r)^q$  as necessary, we obtain that  $A$  is a regular matrix with the entries having the required form.

As  $A$  is (can be picked to be) regular, we obtain from (17)

$$A^{-1} \cdot \begin{pmatrix} \vdots \\ v_{2^q}^{\bar{p}, \bar{\tau}_e} \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ S_e \cdot z_{e(\bar{p})} \\ \vdots \end{pmatrix} \tag{19}$$



and we can now obtain  $z_{\vec{x}} = z_{e(\vec{p})}$  as a difference of functions  $v_{2^q}^{\vec{p}, \vec{\tau}_e}$ : we have

$$z_{\vec{x}} = \frac{1}{S_e} \cdot \sum_{e \in E} b_e \cdot v_{2^q}^{\vec{p}, \vec{\tau}_e}$$

and collecting the  $v_{2^q}^{\vec{p}, \vec{\tau}_e}$  with positive coefficients in one term, we obtain

$$z_{\vec{x}} = \gamma \cdot w_1 - \lambda \cdot w_2, \tag{20}$$

where  $w_1$  is a probability function consisting of the  $v_{2^q}^{\vec{p}, \vec{\tau}_e}$  with positive coefficients and  $w_2$  a probability function consisting of the  $v_{2^q}^{\vec{p}, \vec{\tau}_e}$  with negative coefficients.

Since  $z_{\vec{x}}(\top) = w_1(\top) = w_2(\top) = 1$ , we must have  $\gamma - \lambda = 1$ , and thus we obtain  $\gamma = 1 + \lambda$ . As furthermore  $w_1, w_2$  are convex combinations of probability functions satisfying ULi with SPx, we have the desired representation for  $z_{\vec{x}}$ . Note that  $\lambda$  depends only on  $2^q$  and the entries of  $A^{-1}$ , which in turn only depend on the  $\vec{\tau}_e$ . As  $z_{\vec{x}}$  only determines  $\vec{p}$ , we obtain a uniform  $\lambda$  for all  $z_{\vec{x}}$ .

Now returning to the general case of  $w$  of the form (14), we obtain

$$\begin{aligned} w &= \int_{\vec{x}} z_{\vec{x}} d\mu(\vec{x}) \\ &= \int_{\vec{x}} (1 + \lambda)w_{1, \vec{x}} - \lambda w_{2, \vec{x}} d\mu(\vec{x}) \\ &= (1 + \lambda) \int_{\vec{x}} w_{1, \vec{x}} d\mu(\vec{x}) - \lambda \int_{\vec{x}} w_{2, \vec{x}} d\mu(\vec{x}) \\ &= (1 + \lambda)w_1 - \lambda w_2 \end{aligned}$$

for some probability functions  $w_1, w_2$  satisfying ULi with SPx, as they are convex combinations of functions satisfying these principles.  $\dashv$

## 5 Conclusion

In this paper we have introduced the principle of Strong Predicate Exchangeability, for which we have provided two de Finetti – style representation theorems. While the principle arose rather by chance, and is more indirectly justified as a rational principle for rational agents, the representation theorems presented in this paper fit nicely into the story of Pure Inductive Logic so far.

Comparing the representation theorems provided in this paper with similar results for Unary Language Invariance it appears there is a recurring theme: Looking at presently known results the building blocks for representing functions satisfying ULi all share the property of Weak Irrelevance. At the same time, there exist representation theorems for

general probability functions satisfying  $\mathcal{P}$  showing these to be *differences* of functions that satisfy ULi with  $\mathcal{P}$ .

As these principles have so far been based on a symmetry based on the language involved, one might be inclined to expect the same behaviour for other principles based on symmetry, at least where it concerns purely unary languages. In terms of polyadic languages, there exists a similar result for Spectrum Exchangeability (generalizing Ax). As the building blocks for the principle SPx discussed in this paper are quite similar to the (unary)  $u^{\bar{p}}$  functions, whose polyadic versions play a role in the representation theorem for Spectrum Exchangeability, one might expect a similar result for a polyadic version of Px, SPx.

## References

- [1] Carnap, R., An Axiom System for Inductive Logic. In P.A. Schilpp, editor, *The Philosophy of Rudolf Carnap*, pages 973–979. Open Court Publishing Company, La Salle, Ill., 1963.
- [2] Cutland, N.J., Loeb measure theory, in *Developments in Nonstandard Mathematics*, Eds. N.J.Cutland, F.Oliveira, V.Neves, J.Sousa-Pinto, Pitman Research Notes in Mathematics Series Vol. 336, Longman, Harlow, 1995, pp151-177.
- [3] Gaifman, H., Concerning measures on first order calculi, *Israel Journal of Mathematics*, 1964, **2**:1-18.
- [4] Kließ, Malte S., *The Principle of Predicate Exchangeability in Pure Inductive Logic*, PhD Thesis, The University of Manchester, Manchester, UK, 2014. Available at <https://www.escholar.manchester.ac.uk/uk-ac-man-scw:219317>
- [5] Kließ, M.S. and Paris, J.B., Predicate Exchangeability and Language Invariance in Pure Inductive Logic, *Logique & Analyse*, 2014, **228**:513-540.
- [6] Paris, J.B. and Vencovská, A., *Pure Inductive Logic*, Perspectives in Logic, Cambridge University Press, 2015.